

Connection established between all 8 nodes and getting total size and we can see all VCPUs adding up in our observation window

```
ubuntu@ip-172-31-46-119: ~/data
File Edit View Search Terminal Help
-rw-r--r-- 1 ubuntu supergroup 100000000 2017-12-04 09:56 /input/in
ubuntu@ip-172-31-46-119:~/data$ hdfs dfs -df
17/12/04 09:56:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin
-java classes where applicable
Filesystem                                         Size        Used      Available  Use%
hdfs://ec2-13-58-159-178.us-east-2.compute.amazonaws.com:9000  467410722816  100805835  342188466176   0%
ubuntu@ip-172-31-46-119:~/data$
```

The screenshot shows the Hadoop Application Overview page for application ID 1512388123747_0002. Key details include:

- User:** ubuntu
- Name:** TeraSort
- Application Type:** MAPREDUCE
- Application Priority:** 0 (Higher Integer value indicates higher priority)
- Yarn Application State:** RUNNING; AM has registered with RM and started running.
- Final Status Reported by AM:** Application has not completed yet.
- Started:** Mon Dec 04 18:02:03 +0000 2017
- Tracking URL:** ApplicationMaster
- Log Aggregation Status:** DISABLED
- Logs:** 0
- Unmanaged Applications:** false
- Application Node Label expression:** <Not set>
- AM container Node Label expression:** <DEFAULT_PARTITION>

Application Metrics:

- Total Resource Preempted: <nmemory:0, vCores:0>
- Total Number of Non-AM Containers Preempted: 0
- Total Number of AM Containers Preempted: 0
- Resource Preemption Action Taken: <nmemory:0, vCores:0>
- Number of Non-AM Containers Preempted from Current Attempt: <nmemory:0, vCores:0>
- Aggregate Resource Allocation: 0 MB-seconds, 011 vcore-seconds
- Aggregate Preempted Resource Allocation: 0 MB-seconds, 0 vcore-seconds

Attempts:

Attempt ID	Started	Node	Logs	Nodes blacklisted by the app	Nodes blacklisted by the system
attempt_1512388123747_0002_2000001	Mon Dec 4 12:02:03 -0600 2017	http://ip-172-31-46-119.us-east-2.compute.internal:8542	0	0	0

For small data On multi node

```
ubuntu@ip-172-31-46-119: ~
File Edit View Search Terminal Help
17/12/03 20:01:00 INFO mapreduce.Job: map 100% reduce 94%
17/12/03 20:01:12 INFO mapreduce.Job: map 100% reduce 100%
17/12/03 20:01:12 INFO mapreduce.Job: Counters: 50
File System Counter
  FILE: Number of bytes read=2062408188
  FILE: Number of bytes written=3995139536
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=1280001428
  HDFS: Number of bytes written=1280000000
  HDFS: Number of small write operations=0
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counter
  Killed map tasks=1
  Launched map tasks=11
  Launched reduce tasks=1
  Data-local map tasks=11
  Total time spent by all maps in occupied slots (ms)=444101
  Total time spent by all reduces in occupied slots (ms)=59743
  Total time spent by all map tasks=444101
  Total time spent by all reduce tasks (ms)=59743
  Total vcore-milliseconds taken by all map tasks=44101
  Total vcore-milliseconds taken by all reduce tasks=59743
  Total megabyte-milliseconds taken by all map tasks=547595424
  Total megabyte-milliseconds taken by all reduce tasks=6116832
Map-Reduce Framework
  Map input records=12800000
  Map output records=12800000
  Map output bytes=1331200060
  Input split bytes=1428
  Combine input records=0
  Combiner output bytes=0
  Reduce input groups=12800000
  Reduce shuffle bytes=1331200060
  Reducer input records=12800000
  Reduce output records=12800000
  Spilled Records=38400000
  Shuffled Maps =10
  Failed Shuffles=0
  Merged Map outputs=10
  Merged Map outputs=10
  GC time elapsed (ms)=1053
  CPU time spent (ms)=114588
  Physical memory (bytes) snapshot=2908971008
  Virtual memory (bytes) snapshot=215045472
  Total committed heap usage (bytes)=2106064896
Shuffle Errors
  BAD_ID=0
  CONNECTION_LOST=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_TYPE=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=12800000000
  File Output Format Counters
  Bytes Written=12800000000
17/12/03 20:01:12 INFO terasort.TeraSort: done
real    2m6.636s
user    0m7.272s
sys     0m0.44s
ubuntu@ip-172-31-46-119: ~
```

SPARK 128 GB Data

```
ubuntu@ip-172-31-11-186: ~
File Edit View Search Terminal Help
ec2-13-58-87-175.us-east-2.compute.amazonaws.com: nodemanager running as process 15343. Stop it first.
ubuntu@ip-172-31-11-186:~/hadoop-2.8.2/sbin$ jps
14692 NameNode
14854 DataNode
17094 Jps
15052 SecondaryNameNode
15343 NodeManager
ubuntu@ip-172-31-11-186:~/hadoop-2.8.2/sbin$ cd
ubuntu@ip-172-31-11-186:~$ spark-shell -i sortSpark.scala
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
17/12/03 10:53:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin
- Java classes where applicable
17/12/03 10:53:43 WARN metastore.ObjectStore: Failed to get database global_temp, returning NoSuchObjectException
Spark context Web UI available at http://172.31.11.186:4040
Spark context available as 'sc' (master = local[*], app id = local-1512298416579).
Spark session available as 'spark'.
Loading sortSpark.scala...
timer_start: Long = 1512298427197
timer_tmp: Long = 1512298427404
file_input: org.apache.spark.rdd.RDD[String] = hdfs://ec2-13-58-87-175.us-east-2.compute.amazonaws.com:9000/input/inp
ut.txt MapPartitionsRDD[1] at textFile at <console>:24
Time taken to take file : 1003In Milliseconds
<console>:25: error: reassignment to val
    timer_tmp=System.currentTimeMillis()
                           ^
file_to_sort: org.apache.spark.rdd.RDD[(String, String)] = MapPartitionsRDD[2] at map at <console>:26
Time taken to map file : -2126In Milliseconds
<console>:25: error: reassignment to val
    timer_tmp=System.currentTimeMillis()
                           ^
[Stage 0:====> (28 + 2) / 477]
```

```
ubuntu@ip-172-31-11-186: ~
File Edit View Search Terminal Help
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
17/12/03 22:48:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
17/12/03 22:48:59 WARN metastore.ObjectStore: Failed to get database global_temp, returning NoSuchObjectException
Spark context Web UI available at http://172.31.11.186:4040
Spark context available as 'sc' (master = local[*], app id = local-1512341323247).
Spark session available as 'spark'.
Loading sortSpark.scala...
timer_start: Long = 1512341333420
file_input: org.apache.spark.rdd.RDD[String] = hdfs://ec2-13-58-87-175.us-east-2.compute.amazonaws.com:9000/input/input.txt MapPartitionsRDD[1] at textFile at <console>:24
file_to_sort: org.apache.spark.rdd.RDD[(String, String)] = MapPartitionsRDD[2] at map at <console>:26
sort_by_key: org.apache.spark.rdd.RDD[(String, org.apache.spark.PartitionedRDD[(String, String)])] at sortByKey at <console>:28
lines: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[3] at map at <console>:29
timer_tmp: Long = 1512352286248
total Time taken to sort file and save :-10954828In Milliseconds
data sorted
Welcome to
version 2.2.0
Using Scala version 2.11.8 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_151)
Type in expressions to have them evaluated.
Type :help for more information.
scala: hdfs
<console>:24: error: not found: value hdfs
      hdfs
      ^
scala: [
```

Spark shell - Environment - Mozilla Firefox

EC2 Manager | EC2 Manager | FINISHED Application | (no subject) - pr | Hadoop Setting | Spark shell - Env | imux cheathelp | Sort Benchmark | apache pig - Me | Hands-on Spark | Amazon EC2 Se | modify-instance | Print the next X | myneweshadoop | +

(ec2-13-58-87-175.us-east-2.compute.amazonaws.com:4040/environment/)

Spark 2.2.0 Jobs Stages Storage Environment Executors SQL

Spark shell application UI

Environment

Runtime Information

Name	Value
Java Home	/usr/lib/jvm/java-8-oracle/jre
Java Version	1.8.0_151 (Oracle Corporation)
Scala Version	version 2.11.8

Spark Properties

Name	Value
spark.app.id	local-1512311740164
spark.app.name	Spark shell
spark.driver.host	172.31.11.186
spark.driver.port	37684
spark.executor.id	driver
spark.home	/usr/local/spark
spark.jars	
spark.master	local[""]
spark.repl.class.outputDir	/home/ubuntu/tmp/spark-18ds310e-fa7-4c9d-87cf-e0514903fd4/repl-749a915a-a9f4-48fa-8564-aa19acea434
spark.repl.class.uri	spark://172.31.11.186:37684/classes
spark.scheduler.mode	FIFO
spark.sql.catalogImplementation	hive
spark.submit.deployMode	client

System Properties

Name	Value
SPARK_SUBMIT	true
awt.toolkit	sun.awt.X11.XToolkit
file.encoding	UTF-8
file.encoding.pkg	sun.io
file.separator	/

Spark shell - Spark jobs - Mozilla Firefox

EC2 Manager | EC2 Manager | FINISHED Application | (no subject) - pr | Hadoop Setting | Spark shell - Sp | imux cheathelp | Sort Benchmark | apache pig - Me | Hands-on Spark | Amazon EC2 Se | modify-instance | Print the next X | myneweshadoop | +

(ec2-13-58-87-175.us-east-2.compute.amazonaws.com:4040/jobs/)

Spark 2.2.0 Jobs Stages Storage Environment Executors SQL

Spark shell application UI

Spark Jobs (?)

User: ubuntu
Total Uptime: 3.2 h
Scheduling Mode: FIFO
Completed Jobs: 2

Event Timeline

Completed Jobs (2)

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
1	saveAsTextFile at <console>:33	2017/12/03 15:16:19	2.4 h	2/2	654/654
0	sortByKey at <console>:28	2017/12/03 14:35:51	40 min	1/1	477/477

Spark shell - Stages for All Jobs - Mozilla Firefox

EC2 Manager | EC2 Management | FINISHED Application | (no subject) - pr | Hadoop: Setting | Spark shell - Stages | imux cheatsheet | Sort Benchmark | apache pig - Mer | Hands-on Spark | Amazon EC2 Service | modify-instance | Print the next X | myneweshadoop | +

(ec2-13-58-87-175.us-east-2.compute.amazonaws.com:4040/stages/)

Search

Spark 2.2.0 Jobs Stages Storage Environment Executors SQL

Spark shell application UI

Stages for All Jobs

Completed Stages: 3

Stage Id	Description	Submitted	Duration	Tasks: Succeeded/Total	Input	Output	Shuffle Read	Shuffle Write
2	saveAsTextFile at <console>:33	+details 2017/12/03 15:56:48	1.8 h	477/477	119.2 GB	66.1 GB		
1	map at <console>:26	+details 2017/12/03 15:16:19	40 min	477/477	119.2 GB		66.1 GB	
0	sortByKey at <console>:28	+details 2017/12/03 14:35:51	40 min	477/477	119.2 GB			

```
File Edit View Search Terminal Help
jps: command not found
ubuntu@ip-172-31-11-186:~/dataMount$ jps
2990 NameNode
3110 DataNode
3384 SecondaryNameNode
7033 Jps
ubuntu@ip-172-31-11-186:~/dataMount$ cd
ubuntu@ip-172-31-11-186:~$ screen spark-shell -i sortSpark.scala
[1] 1188 [detached]
ubuntu@ip-172-31-11-186:~$ hdfs dfs -ls /output
17/12/04 02:09:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 478 items
-rw-r--r-- 1 ubuntu supergroup 0 2017-12-04 01:51 /output/_SUCCESS
-rw-r--r-- 1 ubuntu supergroup 282688580 2017-12-04 00:06 /output/part-00000
-rw-r--r-- 1 ubuntu supergroup 271557700 2017-12-04 00:06 /output/part-00001
-rw-r--r-- 1 ubuntu supergroup 211316380 2017-12-04 00:06 /output/part-00002
-rw-r--r-- 1 ubuntu supergroup 269946280 2017-12-04 00:07 /output/part-00003
-rw-r--r-- 1 ubuntu supergroup 359934880 2017-12-04 00:07 /output/part-00004
-rw-r--r-- 1 ubuntu supergroup 351734880 2017-12-04 00:07 /output/part-00005
-rw-r--r-- 1 ubuntu supergroup 303108380 2017-12-04 00:08 /output/part-00006
-rw-r--r-- 1 ubuntu supergroup 228930680 2017-12-04 00:08 /output/part-00007
-rw-r--r-- 1 ubuntu supergroup 295661580 2017-12-04 00:08 /output/part-00008
-rw-r--r-- 1 ubuntu supergroup 2811316380 2017-12-04 00:08 /output/part-00010
-rw-r--r-- 1 ubuntu supergroup 2910584280 2017-12-04 00:09 /output/part-00011
-rw-r--r-- 1 ubuntu supergroup 285561880 2017-12-04 00:09 /output/part-00013
-rw-r--r-- 1 ubuntu supergroup 302776580 2017-12-04 00:09 /output/part-00014
-rw-r--r-- 1 ubuntu supergroup 256829780 2017-12-04 00:10 /output/part-00015
-rw-r--r-- 1 ubuntu supergroup 244348880 2017-12-04 00:10 /output/part-00017
-rw-r--r-- 1 ubuntu supergroup 213327880 2017-12-04 00:10 /output/part-00019
-rw-r--r-- 1 ubuntu supergroup 243773180 2017-12-04 00:11 /output/part-00020
-rw-r--r-- 1 ubuntu supergroup 274084780 2017-12-04 00:11 /output/part-00022
-rw-r--r-- 1 ubuntu supergroup 314705580 2017-12-04 00:11 /output/part-00023
-rw-r--r-- 1 ubuntu supergroup 2910584280 2017-12-04 00:12 /output/part-00024
-rw-r--r-- 1 ubuntu supergroup 239630880 2017-12-04 00:12 /output/part-00025
-rw-r--r-- 1 ubuntu supergroup 215629980 2017-12-04 00:12 /output/part-00026
-rw-r--r-- 1 ubuntu supergroup 256393480 2017-12-04 00:12 /output/part-00027
-rw-r--r-- 1 ubuntu supergroup 249474580 2017-12-04 00:13 /output/part-00029
-rw-r--r-- 1 ubuntu supergroup 278755380 2017-12-04 00:13 /output/part-00030
-rw-r--r-- 1 ubuntu supergroup 272803880 2017-12-04 00:13 /output/part-00031
-rw-r--r-- 1 ubuntu supergroup 272803880 2017-12-04 00:13 /output/part-00032
-rw-r--r-- 1 ubuntu supergroup 278188180 2017-12-04 00:14 /output/part-00033
-rw-r--r-- 1 ubuntu supergroup 280188180 2017-12-04 00:14 /output/part-00034
-rw-r--r-- 1 ubuntu supergroup 243828180 2017-12-04 00:14 /output/part-00035
-rw-r--r-- 1 ubuntu supergroup 368729980 2017-12-04 00:15 /output/part-00036
-rw-r--r-- 1 ubuntu supergroup 257310780 2017-12-04 00:14 /output/part-00037
-rw-r--r-- 1 ubuntu supergroup 247407480 2017-12-04 00:15 /output/part-00038
-rw-r--r-- 1 ubuntu supergroup 274382380 2017-12-04 00:15 /output/part-00039
-rw-r--r-- 1 ubuntu supergroup 244669280 2017-12-04 00:15 /output/part-00040
-rw-r--r-- 1 ubuntu supergroup 241883280 2017-12-04 00:16 /output/part-00041
-rw-r--r-- 1 ubuntu supergroup 241883280 2017-12-04 00:16 /output/part-00042
-rw-r--r-- 1 ubuntu supergroup 268076580 2017-12-04 00:16 /output/part-00043
-rw-r--r-- 1 ubuntu supergroup 292833580 2017-12-04 00:16 /output/part-00044
-rw-r--r-- 1 ubuntu supergroup 226916080 2017-12-04 00:16 /output/part-00046
-rw-r--r-- 1 ubuntu supergroup 268335980 2017-12-04 00:17 /output/part-00047
-rw-r--r-- 1 ubuntu supergroup 368556980 2017-12-04 00:17 /output/part-00049
-rw-r--r-- 1 ubuntu supergroup 289515680 2017-12-04 00:17 /output/part-00050
-rw-r--r-- 1 ubuntu supergroup 261061680 2017-12-04 00:18 /output/part-00051
-rw-r--r-- 1 ubuntu supergroup 298249080 2017-12-04 00:18 /output/part-00052
```

SPARK 1TB Processing:

Spark shell - Stages for All Jobs - Mozilla Firefox

Billing Management Conso EC2 Management Console Spark shell - Stages for All Jobs Apache Spark Quick Guide Sent Mail - prath1@hawk.i Problem loading page PA2_report - Google Docs Screenshots - Google Docs

ec2-18-217-198-64.us-east-2.compute.amazonaws.com:4040/stages/

Spark 2.2.0 Jobs Stages Storage Environment Executors SQL

Spark shell application UI

Stages for All Jobs

Completed Stages: 3

Completed Stages (3)

Stage Id	Description	Submitted	Duration	Tasks: Succeeded/Total	Input	Output	Shuffle Read	Shuffle Write
2	saveAsTextFile at <console>:33	+details 2017/12/04 01:29:57	7.3 h	1863/1863	931.3 GB	523.5 GB		
1	map at <console>:26	+details 2017/12/04 00:21:08	1.1 h	1863/1863	931.4 GB		523.5 GB	
0	sortByKey at <console>:28	+details 2017/12/03 23:48:23	33 min	1863/1863	931.4 GB			

Spark shell - Details for Stage 2 (Attempt 0) - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Spark shell - Details for Stage 2 +

ec2-13-58-87-175.us-east-2.compute.amazonaws.com:4040/stages/stage/?id=2&attempt=0

Spark 2.2.0 Jobs Stages Storage Environment Executors SQL

Spark shell application UI

Details for Stage 2 (Attempt 0)

Total Time Across All Tasks: 3.5 h
 Locality Level Summary: Any 477
 Output: 119.2 GB / 1280000000
 Shuffle Read: 66.1 GB / 1280000000
 Shuffle Spill (Memory): 309.1 GB
 Shuffle Spill (Disk): 59.2 GB

▶ DAG Visualization
 ▶ Show Additional Metrics
 ▶ Event Timeline

Summary Metrics for 477 Completed Tasks

Metric	Min	25th percentile	Median	75th percentile	Max
Duration	16 s	23 s	26 s	29 s	49 s
GC Time	4 s	8 s	10 s	12 s	22 s
Output Size / Records	170.2 MB / 1784218	232.8 MB / 2440692	254.4 MB / 2667139	279.0 MB / 2925480	420.3 MB / 4407553
Shuffle Read Size / Records	94.6 MB / 1784218	129.2 MB / 2440692	141.1 MB / 2667139	154.7 MB / 2925480	232.8 MB / 4407553
Shuffle spill (memory)	375.9 MB	563.9 MB	737.5 MB	751.9 MB	1127.8 MB
Shuffle spill (disk)	72.4 MB	108.8 MB	133.9 MB	144.8 MB	217.4 MB

Aggregated Metrics by Executor

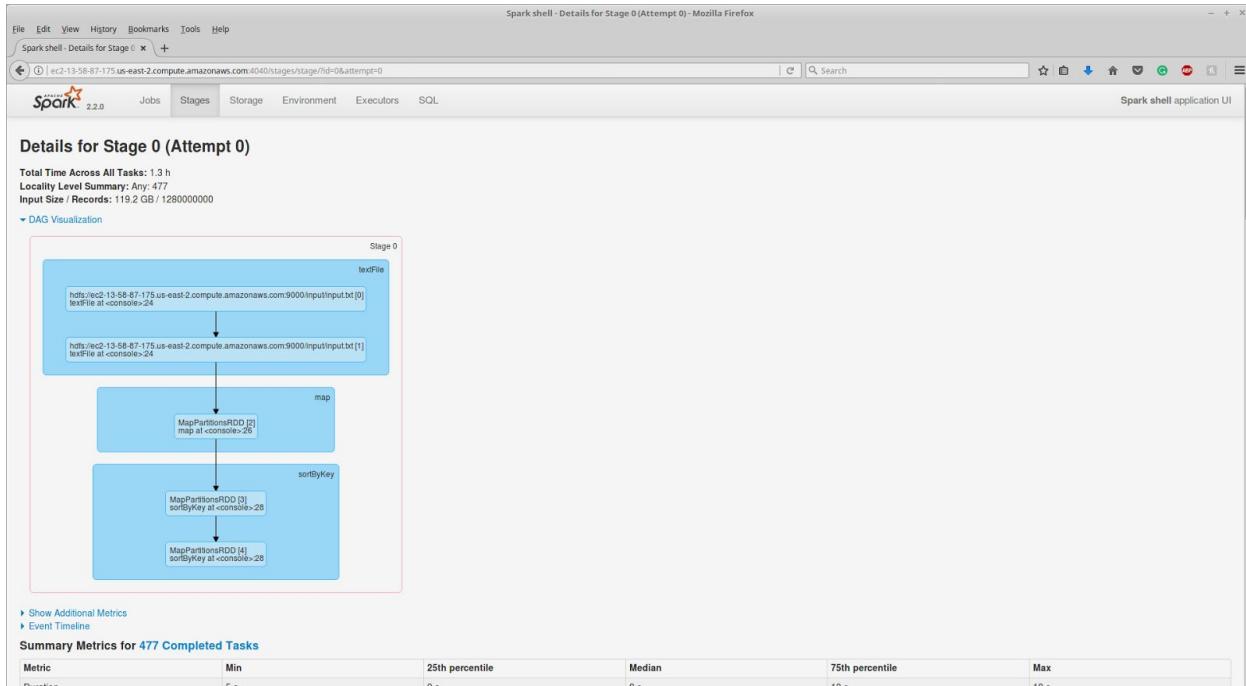
Executor ID	Address	Task Time	Total Tasks	Failed Tasks	Killed Tasks	Succeeded Tasks	Output Size / Records	Shuffle Read Size / Records	Shuffle Spill (Memory)	Shuffle Spill (Disk)	Blacklisted
driver	172.31.11.186:36189	3.5 h	477	0	0	477	119.2 GB / 1280000000	66.1 GB / 1280000000	309.1 GB	59.2 GB	0

Tasks (477)

Page: 1 2 3 4 5 >

5 Pages, Jump to 1 Show 100 items in a page Go

Index	ID	Attempt	Status	Locality Level	Executor ID / Host	Launch Time	Duration	GC Time	Output Size / Records	Shuffle Read Size / Records	Shuffle Spill (Memory)	Shuffle Spill (Disk)	Errors
0	954	0	SUCCESS	ANY	driver / localhost	2017/12/04 00:06:07	32 s	7 s	269.6 MB / 2826885	149.5 MB / 2826885	751.9 MB	144.9 MB	
1	955	0	SUCCESS	ANY	driver / localhost	2017/12/04 00:06:07	31 s	7 s	261.7 MB / 2744557	145.2 MB / 2744557	751.9 MB	145.0 MB	
2	956	0	SUCCESS	ANY	driver / localhost	2017/12/04 00:06:38	30 s	10 s	297.0 MB / 3119785	164.6 MB / 3119785	751.9 MB	144.3 MB	
3	957	0	SUCCESS	ANY	driver / localhost	2017/12/04 00:06:39	27 s	10 s	257.4 MB / 2699462	142.8 MB / 2699462	751.9 MB	145.0 MB	
4	958	0	SUCCESS	ANY	driver / localhost	2017/12/04 00:07:06	74 s	14 s	218.1 MB / 3205348	176.4 MB / 3205348	751.9 MB	144.4 MB	



Result

```

File Edit View Search Terminal Help
Checksum: command not found
adoop: command not found
SUCCESS: command not found
ubuntu@ip-172-31-11-186:~$ 
cat: part-*: No such file or directory
ubuntu@ip-172-31-11-186:~$ ls
32_code  ec2-18-217-198-64.us-east-2.compute.amazonaws.com  hadoop-2.8.2  Holmes.txt  metastore_db  parallel-ssh  spark-2.2.0-bin-hadoop2.7.tgz  TeraSort.jar
64      bigfile derby.log  gpl-2.0.txt  hadoopptp  input  output  sortSpark.scala  TeraSort.Hadoop.java  Terav1.java
ubuntu@ip-172-31-11-186:~$ cd dataMount/
ubuntu@ip-172-31-11-186:~/dataMount$ ls
hdfs part-00024 part-00050 part-00076 part-00102 part-00128 part-00154 part-00180 part-00206 part-00232 part-00258 part-00284 part-00310 part-00336 part-00362 part-00388 part-00414 part-00440 part-00466
lost+found part-00025 part-00051 part-00077 part-00103 part-00129 part-00155 part-00181 part-00207 part-00233 part-00259 part-00285 part-00311 part-00337 part-00363 part-00389 part-00415 part-00441 part-00467
part-00009 part-00026 part-00052 part-00078 part-00104 part-00130 part-00156 part-00182 part-00208 part-00234 part-00260 part-00286 part-00312 part-00338 part-00364 part-00390 part-00416 part-00442 part-00468
part-00091 part-00027 part-00053 part-00079 part-00105 part-00131 part-00157 part-00183 part-00209 part-00235 part-00261 part-00287 part-00313 part-00339 part-00365 part-00391 part-00417 part-00443 part-00469
part-00002 part-00028 part-00054 part-00080 part-00106 part-00132 part-00158 part-00184 part-00210 part-00238 part-00262 part-00288 part-00314 part-00340 part-00366 part-00392 part-00418 part-00444 part-00470
part-00003 part-00029 part-00055 part-00081 part-00107 part-00133 part-00160 part-00185 part-00211 part-00239 part-00263 part-00289 part-00315 part-00341 part-00367 part-00393 part-00419 part-00445 part-00471
part-00004 part-00030 part-00056 part-00082 part-00108 part-00134 part-00161 part-00186 part-00212 part-00240 part-00268 part-00296 part-00316 part-00342 part-00368 part-00394 part-00420 part-00446 part-00472
part-00005 part-00031 part-00057 part-00083 part-00109 part-00135 part-00162 part-00187 part-00213 part-00241 part-00265 part-00291 part-00317 part-00343 part-00369 part-00395 part-00421 part-00447 part-00473
part-00006 part-00032 part-00058 part-00084 part-00110 part-00136 part-00163 part-00191 part-00214 part-00242 part-00266 part-00292 part-00318 part-00345 part-00370 part-00396 part-00422 part-00448 part-00474
part-00007 part-00033 part-00059 part-00085 part-00111 part-00137 part-00164 part-00192 part-00215 part-00243 part-00267 part-00293 part-00319 part-00346 part-00371 part-00397 part-00423 part-00450 part-00475
part-00008 part-00034 part-00060 part-00086 part-00112 part-00138 part-00165 part-00193 part-00216 part-00244 part-00268 part-00294 part-00320 part-00345 part-00372 part-00398 part-00424 part-00458 part-00476
part-00009 part-00035 part-00061 part-00087 part-00113 part-00139 part-00166 part-00194 part-00217 part-00245 part-00269 part-00295 part-00321 part-00347 part-00373 part-00399 part-00425 part-00451 part-SUCCESS
part-00010 part-00036 part-00062 part-00088 part-00114 part-00140 part-00167 part-00195 part-00219 part-00246 part-00271 part-00297 part-00323 part-00349 part-00375 part-00399 part-00426 part-00452 part-SUCCESS
part-00011 part-00037 part-00063 part-00089 part-00115 part-00141 part-00168 part-00196 part-00220 part-00247 part-00272 part-00298 part-00324 part-00350 part-00376 part-00402 part-00453 part-SUCCESS
part-00012 part-00038 part-00064 part-00090 part-00116 part-00142 part-00169 part-00197 part-00221 part-00248 part-00273 part-00299 part-00325 part-00351 part-00377 part-00403 part-00454 part-SUCCESS
part-00013 part-00039 part-00065 part-00091 part-00117 part-00143 part-00170 part-00198 part-00222 part-00249 part-00274 part-00300 part-00326 part-00352 part-00378 part-00404 part-00455 part-SUCCESS
part-00014 part-00040 part-00066 part-00092 part-00118 part-00144 part-00171 part-00199 part-00223 part-00250 part-00275 part-00301 part-00327 part-00353 part-00379 part-00405 part-00456 part-SUCCESS
part-00015 part-00041 part-00067 part-00093 part-00119 part-00145 part-00172 part-00200 part-00251 part-00276 part-00302 part-00328 part-00354 part-00381 part-00406 part-00457 part-SUCCESS
part-00016 part-00042 part-00068 part-00094 part-00120 part-00146 part-00173 part-00201 part-00252 part-00277 part-00303 part-00330 part-00355 part-00382 part-00407 part-00443 part-00458 part-SUCCESS
part-00017 part-00043 part-00069 part-00095 part-00121 part-00147 part-00174 part-00202 part-00253 part-00278 part-00304 part-00331 part-00356 part-00383 part-00408 part-00444 part-00459 part-SUCCESS
part-00018 part-00044 part-00070 part-00096 part-00122 part-00148 part-00175 part-00203 part-00254 part-00280 part-00305 part-00332 part-00357 part-00384 part-00409 part-00445 part-00460 part-SUCCESS
part-00019 part-00045 part-00071 part-00097 part-00123 part-00149 part-00176 part-00204 part-00255 part-00281 part-00306 part-00333 part-00358 part-00385 part-00410 part-00446 part-00461 part-SUCCESS
part-00020 part-00046 part-00072 part-00098 part-00124 part-00150 part-00177 part-00205 part-00256 part-00282 part-00307 part-00334 part-00359 part-00386 part-00411 part-00447 part-00462 part-SUCCESS
part-00021 part-00047 part-00073 part-00099 part-00125 part-00151 part-00178 part-00206 part-00257 part-00283 part-00308 part-00335 part-00360 part-00387 part-00412 part-00448 part-00463 part-SUCCESS
part-00022 part-00048 part-00074 part-00100 part-00126 part-00152 part-00179 part-00207 part-00258 part-00284 part-00309 part-00336 part-00361 part-00388 part-00413 part-00449 part-00464 part-SUCCESS
part-00023 part-00049 part-00075 part-00101 part-00127 part-00153 part-00180 part-00208 part-00259 part-00285 part-00310 part-00337 part-00362 part-00391 part-00414 part-00450 part-00465 part-SUCCESS
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort dataMount/part-00000
[bash] ./4valsort: No such file or directory
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort part-00000
[bash] ./4valsort: No such file or directory
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort part-00000
Records: 159175216711d9
Duplicate keys: 0
SUCCESS: all records are in order
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort part-00032
Records: 278088
Checksum: 14cf7f8593377f
Records: 0
SUCCESS: all records are in order
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort part-00023
sp_start failed: part-000283: No such file or directory
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort part-00023
Records: 0
Checksum: 17454637a8af6b
Records: 0
SUCCESS: all records are in order
ubuntu@ip-172-31-11-186:~/dataMount$ ./4valsort part-00213
Records: 3657198
Checksum: 1754637a8af6b
Records: 0
SUCCESS: all records are in order
ubuntu@ip-172-31-11-186:~/dataMount$ 

```

HADOOP 128 GB Data

Instances

Tags Reports Limits

Instances

Launch Templates Spot Requests

Instance State: Running

Add filter

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6 IPs	Key Name	Monitoring
Hadoop 1TB	i-0072d140647478zf	i3.4xlarge	us-east-2c	running	2/2 checks ...	None	ec2-18-217-198-64.us-e...	18.217.198.64	-	pk_First	disabled
Hadoop-Main-singleNode	i-096a9b9ceec1d574a	i3.large	us-east-2a	running	2/2 checks ...	None	ec2-18-216-66-98.us-ea...	18.216.66.98	-	pk_First	disabled

1 to 2 of 2

EC2 Management Console - Mozilla Firefox

https://us-east-2.console.aws.amazon.com/ec2v2/home?region=us-east-2#instances?sort=instanceState

Services Resource Groups

Launch Instance Connect Actions

Filter by tags and attributes or search by keyword

1 to 17 of 17

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6 IPs	Key Name	Monitoring
DataNode1	i-0118903cfc03ded3fb	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-13-59-181-1.us-eas...	13.59.181.1	-	pk_First	disabled
DataNode2	i-02bcfd72f699ed24c	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-13-59-55-233.us.ea...	13.59.55.233	-	pk_First	disabled
DataNode3	i-03ca26349f729e0b4	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-221-22-244.us.e...	18.221.22.244	-	pk_First	disabled
DataNode4	i-03de6090e0238f43	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-217-170-109.us...	18.217.170.109	-	pk_First	disabled
DataNode5	i-0973884b9058bd83	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-38-67-172.us.ea...	18.38.67.172	-	pk_First	disabled
DataNode6	i-0eb684ca363cf37cb	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-217-181-33.us.e...	18.217.181.33	-	pk_First	disabled
DataNode7	i-0fafc49fc19645f2	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-52-18-248-69.us.ea...	52.18.248.69	-	pk_First	disabled
Hadoop 1TB	i-00720140547478bf	i3.4xlarge	us-east-2c	running	2/2 checks ...	None	ec2-18-217-198-64.us.e...	18.217.198.64	-	pk_First	disabled
ARchiExtSort	i-082a26ca4b12ceb20	i3.large	us-east-2a	running	2/2 checks ...	None	ec2-18-216-141-212.us...	18.216.141.212	-	pk_First	disabled
Hadoop-Main-singleNode	i-096a3b9ceec1d574a	i3.large	us-east-2a	running	2/2 checks ...	None	ec2-13-58-87-175.us.east-2.compute.amazonaws.com	13.58.87.175	-	pk_First	disabled
hadoop-dataNode1	i-007ffea21a427290	t2.micro	us-east-2c	stopped	0/2 checks ...	None	-	-	-	pk_First	disabled
hadoop-dataNode2	i-0284000009025e5d	t2.micro	us-east-2b	stopped	0/2 checks ...	None	-	-	-	pk_First	disabled
hadoop-dataNode3	i-05c2a189a20c21ee	t2.micro	us-east-2b	stopped	0/2 checks ...	None	-	-	-	pk_First	disabled
1Node-hadoop	i-0ce59fe2cc7e3db	i3.large	us-east-2b	stopped	0/2 checks ...	None	-	-	-	pk_First	disabled
hadoop-NameNode	i-095e252d3afeb1b2	t2.micro	us-east-2b	stopped	0/2 checks ...	None	-	-	-	pk_First	disabled

Feedback English (US)

© 2008-2017, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

```
ubuntu@ip-172-31-11-186:~$ hdfs dfs -ls /output
17/12/02 12:01:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
-rw-r--r-- 1 ubuntu supergroup          0 2017-12-02 11:32 /output/_SUCCESS
-rw-r--r-- 10 ubuntu supergroup         0 2017-12-02 06:55 /output/_partition.lst
-rw-r--r-- 1 ubuntu supergroup 128000000000 2017-12-02 11:32 /output/part-r-00000
ubuntu@ip-172-31-11-186:~$ hdfs dfs -copyToLocal /output/part-r-00000 ~/tmp/
```

Namenode information

ec2-18-216-66-98.us-east-2.compute.amazonaws.com:50070/dfshealth.html#tab-overview

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Overview 'ec2-18-216-66-98.us-east-2.compute.amazonaws.com:9000' (active)

Started:	Sat Dec 02 00:50:56 -0600 2017
Version:	2.8.2, f66c47fa01a19637879e95fb0c41f798373828f6
Compiled:	Thu Oct 19 15:39:00 -0500 2017 by jdu from branch-2.8.2
Cluster ID:	CID-7d44f701-083e-49ac-b857-c7780232d001
Block Pool ID:	BP-872161040-172.31.11.186-1512179248512

Summary

Security is off.
Safemode is off.

18 files and directories, 957 blocks = 975 total filesystem object(s).

Heap Memory used 41.88 MB of 291 MB Heap Memory. Max Heap Memory is 889 MB.

Non Heap Memory used 57.53 MB of 58.84 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	435.31 GB
DFS Used:	240.29 GB (55.2%)
Non DFS Used:	73.66 MB
DFS Remaining:	172.82 GB (39.7%)
Block Pool Used:	240.29 GB (55.2%)
DataNodes usages% (Min/Median/Max/stdDev):	55.20% / 55.20% / 55.20% / 0.00%
Live Nodes	1 (Decommissioned: 0)
Dead Nodes	0 (Decommissioned: 0)

Namenode information - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Namenode information

ec2-18-216-66-98.us-east-2.compute.amazonaws.com:50070/dshealth.html#tab-overview

Dead Nodes

Decommissioning Nodes	0 (0 decommissioned, 0)
Total Datanode Volume Failures	0 (0 B)
Number of Under-Replicated Blocks	0
Number of Blocks Pending Deletion	0
Block Deletion Start Time	Sat Dec 02 00:50:56 -0600 2017
Last Checkpoint Time	Sat Dec 02 05:52:27 -0600 2017

NameNode Journal Status

Current transaction ID: 5175

Journal Manager	State
FileJournalManager(root=/home/ubuntu/dataMount/hdfs/namenode)	EditLogFileOutputStream(/home/ubuntu/dataMount/hdfs/namenode/current/edits_inprogress_0000000000000005175)

NameNode Storage

Storage Directory	Type	State
/home/ubuntu/dataMount/hdfs/namenode	IMAGE_AND_EDITS	Active

DFS Storage Types

Storage Type	Configured Capacity	Capacity Used	Capacity Remaining	Block Pool Used	Nodes In Service
DISK	435.31 GB	240.29 GB (55.2%)	172.82 GB (39.7%)	240.29 GB	1

Application Attempt appattempt_1512197486141_0001_000001 - Mozilla Firefox

Application Attempt appattempt_1512197486141_0001_000001

Logged in as drwho

hadoop

Application Attempt Overview

Application Attempt State: FINISHED
Started: Sat Dec 02 06:55:43 +0000 2017
Elapsed: 4hrs, 36mins, 53sec
AM Container: container_1512197486141_0001_01_000001
Node: ip-172-31-11-186.us-east-2.compute.internal:35848
Tracking URL: History
Diagnostics info:
Nodes blacklisted by the application: -
Nodes blacklisted by the system: -

Total Allocated Containers: 480

Each table cell represents the number of NodeLocal/RackLocal/OffSwitch containers satisfied by NodeLocal/RackLocal/OffSwitch resource requests.

	Node Local Request	Rack Local Request	Off Switch Request
Num Node Local Containers (satisfied by)	478	0	2
Num Rack Local Containers (satisfied by)	0	0	0
Num Off Switch Containers (satisfied by)	0	0	2

Show 20 entries

Container ID	Node	Container Exit Status
No data available in table		

Search: Logs

Showing 0 to 0 of 0 entries

First Previous Next Last

Result:

```
ubuntu@ip-172-31-11-186: ~/tmp
File Edit View Search Terminal Help
ubuntu@ip-172-31-11-186:~$ hdfs dfs -copyToLocal /output/part-r-00000 ~/tmp/
17/12/02 12:02:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
ubuntu@ip-172-31-11-186:~$ ls
32_code derby.log gpl-2.0.txt hadoop-2.8.2.tar.gz Holmes.
txt input newData parallel-ssh TeraSort_Hadoop.java Terasort.java
[...]
64_dataMount gensor.tgz hadoop-2.8.2 hadooppmp init_se
tup.sh metastore_db output q TeraSort.jar Terav1.java
ubuntu@ip-172-31-11-186:~$ cd 64/
ubuntu@ip-172-31-11-186:~/64$ ./
gensor valsor
ubuntu@ip-172-31-11-186:~/64$ ./ www.example.com instead of www.example.com
gensor valsor
ubuntu@ip-172-31-11-186:~/64$ cp valsor ../tmp/valsor
ubuntu@ip-172-31-11-186:~/64$ cd ../tmp/
ubuntu@ip-172-31-11-186:~/tmp$ ls
hadooppmp lost+found part-r-00000 valsor
ubuntu@ip-172-31-11-186:~/tmp$ ./valsor part-r-00000
Records: 1280000000
Checksum: 2625aa9ecff4ff117
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-11-186:~/tmp$
```

HADOOP 1TB:

EC2 Management Console - Mozilla Firefox

EC2 Management Con: EC2 Management Con: RUNNING Applications: raid - prath1@hawk.it: Apache Spark Quick C: Sparkshell - Stages for All: Sort Benchmark Data: apache pig - Merging: Hands-on Spark Tutor: How to Setup an Apac...: zombaki: Ohio: Support: https://us-east-2.console.aws.amazon.com/ec2v2/home?region=us-east-2#instances;sort=instanceState

EC2 Dashboard Services Resource Groups ▾

Launch Instance Connect Actions ▾

Filter by tags and attributes or search by keyword 1 to 17 of 17

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6 IPs	Key Name	Monitoring		
DataNode1	i-0118939c03ded3fb	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-13-59-181-1.us-eas...	13.59.181.1	-	pk_First	disabled		
DataNode2	i-026cf72f699e6d24c	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-13-59-55-233.us.ea...	13.59.55.233	-	pk_First	disabled		
DataNode3	i-03a26149f7129e64	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-221-22-244.us.e...	18.221.22.244	-	pk_First	disabled		
DataNode4	i-03de0609e0238f43	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-217-170.109.us...	18.217.170.109	-	pk_First	disabled		
DataNode5	i-0973884b9508bd83	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-13-58-67-172.us.ea...	13.58.67.172	-	pk_First	disabled		
DataNode6	i-09e684ca363cf37cb	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-217-181-33.us.e...	18.217.181.33	-	pk_First	disabled		
DataNode7	i-0fafc49fc1b945f52	i3.large	us-east-2c	running	2/2 checks ...	None	ec2-18-14-246-69.us.ea...	52.14.246.69	-	pk_First	disabled		
Hadoop 1TB	i-00720d40647478bf	i3.4xlarge	us-east-2c	running	2/2 checks ...	None	ec2-18-217-198-64.us.e...	18.217.198.64	-	pk_First	disabled		
ARchisExtSort	i-080a26ca4b12ce020	i3.large	us-east-2a	running	2/2 checks ...	None	ec2-18-216-141-212.us...	18.216.141.212	-	pk_First	disabled		
Hadoop-Main-singleNode	i-095e3b9cc1c574a	i3.large	us-east-2a	running	2/2 checks ...	None	ec2-13-58-67-175.us.ea...	13.58.67.175	-	pk_First	disabled		
Description	Status Checks	Monitoring	Tags										
				Hadoop 1TB									
Instance ID	i-00720d40647478bf			Public DNS (IPv4): ec2-18-217-198-64.us-east-2.compute.amazonaws.com									
Instance state	running			IPv4 Public IP: 18.217.198.64									
Instance type	i3.4xlarge			IPv6 IPs: -									
Elastic IPs				Private DNS: ip-172-31-35-123.us-east-2.compute.internal									
Availability zone	us-east-2c			Private IPs: 172.31.35.123									
Security groups	launch-wizard-6, view inbound rules			Secondary private IPs: -									
Scheduled events	No scheduled events			VPC ID: vpc-a64550cf									
AMI ID	ubuntu/images/hvm-ssd/ubuntu-xenial-16.04-amd64-server-20171121.1 (ami-82f4daef)			Subnet ID: subnet-b852144c8									
Platform	-			Network interfaces: eth0									

Feedback English (US)

Application application_1512256721841_0003 - Mozilla Firefox

EC2 Manager Application appli: raid - prath1@: How to move fil: Sort Benchmark: apache pig - Me: Hands-on Spark: How to Setup an Amazon EC2: modify-instance: Apache Spark: create map in s: +

Logged in as drwho

Application application_1512256721841_0003

Cluster

- About
- Nodes
- Labels
- Applications
- NEW
- NEW_SAVING
- SUCCEEDED
- ACCEPTED
- RUNNING
- FINISHED
- FAILED
- KILLED
- Scheduler
- Tools

Kill Application

User: ubuntu
Name: TeraSort
Application Type: MAPREDUCE
Application Tags:
Application Priority: 0 (Higher Integer value indicates higher priority)
YarnApplicationState: FINISHED
FinalStatus Reported by AM: SUCCEEDED
Started: Sun Dec 03 01:31:18 +0000 2017
Elapsed: 11hrs, 24mins, 56sec
Tracking URL: History
Log Aggregation Status: DISABLED
Diagnostics:
Unmanaged Application: false
Application Node Label expression: <Not set>
AM container Node Label expression: <DEFAULT_PARTITION>

Application Overview

Total Resource Preempted: <memory-0, vCores:0>
Total Number of Non-AM Containers Preempted: 0
Total Number of AM Containers Preempted: 0
Resource Preempted From Current Attempt: <memory-0, vCores:0>
Number of Non-AM Containers Preempted From Current Attempt: 0
Aggregate Resource Allocation: 184632929 MB-seconds, 138269 vcore-seconds
Aggregate Preempted Resource Allocation: 0 MB-seconds, 0 vcore-seconds

Application Metrics

Show 20 entries

Attempt ID	Started	Node	Logs	Nodes blacklisted by the app	Nodes blacklisted by the system
appattempt_1512256721841_0003_000001	Sat Dec 2 19:31:18 -0600 2017	http://ip-172-31-35-123.us-east-2.compute.internal:8042	0	0	0

Showing 1 to 1 of 1 entries

The screenshot shows the Hadoop Application Overview page for a completed job attempt. The main header reads "Application Attempt appattempt_1512256721841_0003_00001". The status is listed as "FINISHED" with a start time of "Sun Dec 03 01:18 +0000 2017" and an elapsed time of "11hrs, 24mins, 56sec". The AM Container and Tracking Node are both listed as "container_1512256721841_0003_01_00001 ip-172-31-35-123.us-east-2.compute.internal:36125". The "Nodes blacklisted by the application" and "Nodes blacklisted by the system" sections are both empty. Below this, a table provides a breakdown of allocated containers: 1864 Node Local Containers (satisfied by), 0 Rack Local Containers (satisfied by), and 0 Off Switch Containers (satisfied by). A note states that each table cell represents the number of NodeLocal/RackLocal/OffSwitch containers satisfied by NodeLocal/RackLocal/OffSwitch resource requests.

The screenshot shows a terminal window displaying the log output of a completed MapReduce job. The log starts with several INFO messages from the map phase, followed by a message indicating the job completed successfully at 12:56:17. The log then details the execution of the reduce phase, including the number of bytes read, operations, and various metrics for map and reduce tasks. It also shows the shuffle and merge process, including the number of failed and successful shuffle files. The log concludes with a summary of total committed heap usage and a final INFO message from Terasort.

Ps-i had to do screen recording as i was using screen and my data was getting lost.so took screen shot from video.

```

hadoop@ip-172-31-35-123:~/data$ hdfs dfs -ls /output
17/12/03 12:59:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
-rw-r--r-- 1 ubuntu supergroup 0 2017-12-03 12:56 /output/_SUCCESS
-rw-r--r-- 10 ubuntu supergroup 0 2017-12-03 01:31 /output/_partition.lst
-rw-r--r-- 1 ubuntu supergroup 10000000000000 2017-12-03 12:56 /output/part-r-00000

```

Result:

```
ubuntu@ip-172-31-35-123: ~/data
File Edit View Search Terminal Help
ubuntu@ip-172-31-35-123:~/data$ ls
32 gsort-linux-1.5.tar.gz hadooptmp lost+found
64 gpl-2.0.txt hdfs output_data_1TB
ubuntu@ip-172-31-35-123:~/data$ ./64/valsrt output_data_1TB
Records: 100000000000
Checksum: 12a06cd06eeb64b16
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-35-123:~/data$
```

The terminal window shows the execution of the gsort command followed by the ./64/valsrt command. The output indicates that 100000000000 records were processed, with a checksum of 12a06cd06eeb64b16 and 0 duplicate keys. The output is labeled as being in order.

Ec2 Login for Shared memory

```
ubuntu@ip-172-31-31-85: ~
archi@archi-linux:~/UbuntuComputer$ ssh -i "pk_First.pem" ubuntu@ec2-18-217-169-166.us-east-2.compute.amazonaws.com
Welcome to Ubuntu 16.04.3 LTS (GNU/Linux 4.4.0-1041-aws x86_64)

 * Documentation: https://help.ubuntu.com
 * Management: https://landscape.canonical.com
 * Support: https://ubuntu.com/advantage

 Get cloud support with Ubuntu Advantage Cloud Guest:
 http://www.ubuntu.com/business/services/cloud

 0 packages can be updated,
 0 updates are security updates.

Last login: Sat Dec  2 00:29:48 2017 from 194.194.122.41
ubuntu@ip-172-31-31-85:~$
```

The screenshot shows an Ubuntu desktop environment. A terminal window is open, displaying the result of an SSH session. The session logs in as 'ubuntu' via 'pk_First.pem' to an AWS instance. The terminal shows the standard Ubuntu 16.04.3 LTS welcome message, package update information, and the last login details. The desktop interface includes a dock with various application icons.

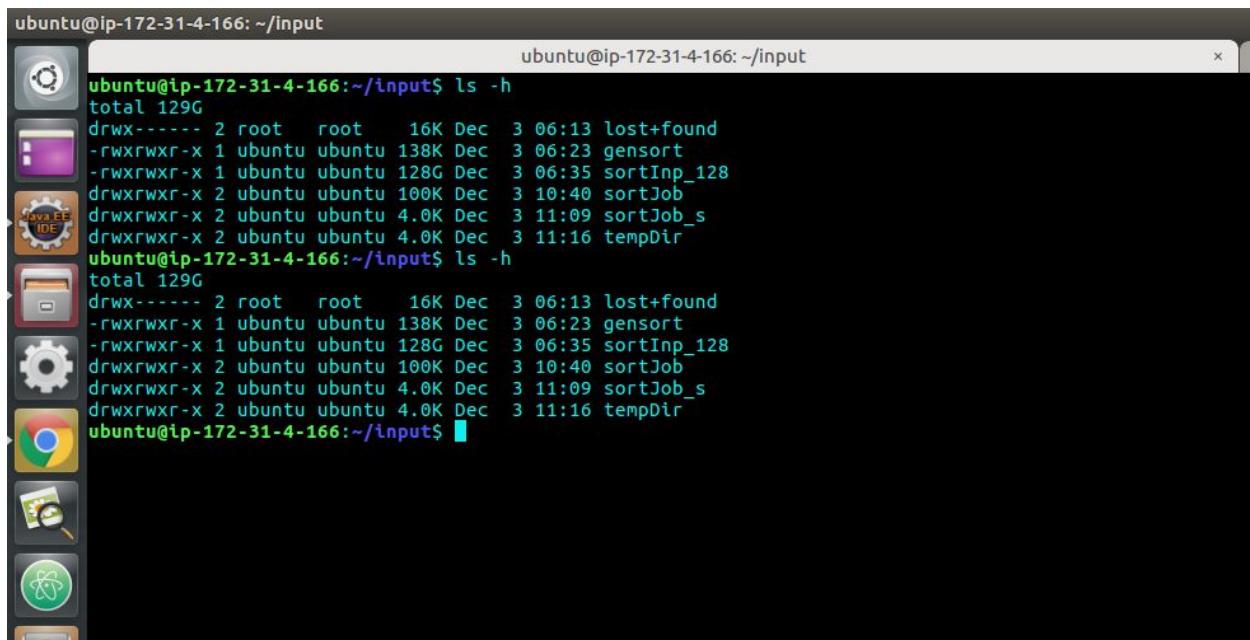
Mounting 2TB EBS volume to EC instance

```
ubuntu@ip-172-31-31-85:~$ java -version
openjdk version "1.8.0_151"
OpenJDK Runtime Environment (build 1.8.0_151-8u151-b12-0ubuntu0.16.04.2-b12)
OpenJDK 64-Bit Server VM (build 25.151-b12, mixed mode)
ubuntu@ip-172-31-31-85:~$ df
Filesystem 1K-blocks Used Available Use% Mounted on
udev 62905972 0 62905972 0% /dev
tmpfs 12582788 8792 12573996 1% /run
/dev/xvda1 8065444 2206592 5842468 28% /
tmpfs 62913936 0 62913936 0% /dev/shm
tmpfs 5120 0 5120 0% /run/lock
tmpfs 62913936 0 62913936 0% /sys/fs/cgroup
tmpfs 12582788 0 12582788 0% /run/user/1000
ubuntu@ip-172-31-31-85:~$ df -h
Filesystem Size Used Avail Use% Mounted on
udev 60G 0 60G 0% /dev
tmpfs 12G 8.6M 12G 1% /run
/dev/xvda1 7.7G 2.2G 5.6G 28% /
tmpfs 60G 0 60G 0% /dev/shm
tmpfs 5.0M 0 5.0M 0% /run/lock
tmpfs 60G 0 60G 0% /sys/fs/cgroup
tmpfs 12G 0 12G 0% /run/user/1000
ubuntu@ip-172-31-31-85:~$ lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
xvda 202:0 0 8G 0 disk
└─xvda1 202:1 0 8G 0 part /
nvme0n1 259:0 0 1.7T 0 disk
nvme0n1 259:1 0 1.7T 0 disk
ubuntu@ip-172-31-31-85:~$ sudo mdadm --create /dev/md0 --run --level=0 --name=RAID --force --raid -devices=2 /dev/nvme0
nvme0 nvme0n1
ubuntu@ip-172-31-31-85:~$ sudo mdadm --create /dev/md0 --run --level=0 --name=RAID --force --raid -devices=2 /dev/nvme0
nvme0 nvme0n1
ubuntu@ip-172-31-31-85:~$ sudo mdadm --create /dev/md0 --run --level=0 --name=RAID --force --raid -devices=2 /dev/nvme0n1 /dev/nvme1n1
mdadm: invalid number of raid devices: -devices=2
ubuntu@ip-172-31-31-85:~$ sudo mdadm --create /dev/md0 --run --level=0 --name=RAID --force --raid-devices=2 /dev/nvme0n1 /dev/nvme1n1
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md0 started.
ubuntu@ip-172-31-31-85:~$ sudo mkfs.ext4 -L RAID /dev/md0
mke2fs 1.42.13 (17-May-2015)
Discarding device blocks: 196612096/927668736
```

Generating 128GB data

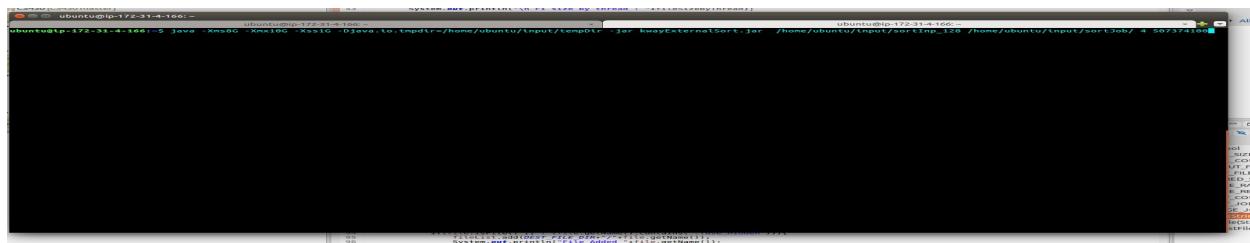
```
ubuntu@ip-172-31-4-166:~/input
ubuntu@ip-172-31-4-166:~/input$ ./gensort -a 1374389534 sortInp_128
```

Confirm Data 128GB



```
ubuntu@ip-172-31-4-166: ~/input
ubuntu@ip-172-31-4-166:~/input$ ls -h
total 129G
drwx----- 2 root root 16K Dec  3 06:13 lost+found
-rwxrwxr-x 1 ubuntu ubuntu 138K Dec  3 06:23 gensort
-rwxrwxr-x 1 ubuntu ubuntu 128G Dec  3 06:35 sortInp_128
drwxrwxr-x 2 ubuntu ubuntu 100K Dec  3 10:40 sortJob
drwxrwxr-x 2 ubuntu ubuntu 4.0K Dec  3 11:09 sortJob_s
drwxrwxr-x 2 ubuntu ubuntu 4.0K Dec  3 11:16 tempDir
ubuntu@ip-172-31-4-166:~/input$ ls -h
total 129G
drwx----- 2 root root 16K Dec  3 06:13 lost+found
-rwxrwxr-x 1 ubuntu ubuntu 138K Dec  3 06:23 gensort
-rwxrwxr-x 1 ubuntu ubuntu 128G Dec  3 06:35 sortInp_128
drwxrwxr-x 2 ubuntu ubuntu 100K Dec  3 10:40 sortJob
drwxrwxr-x 2 ubuntu ubuntu 4.0K Dec  3 11:09 sortJob_s
drwxrwxr-x 2 ubuntu ubuntu 4.0K Dec  3 11:16 tempDir
ubuntu@ip-172-31-4-166:~/input$
```

Running 128 GB



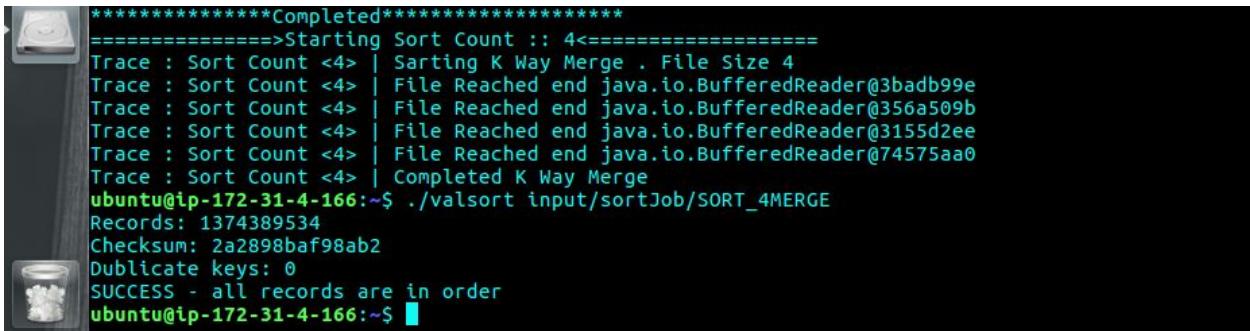
Running confirmation 128 Sort

```
ubuntu@ip-172-31-4-166:~/input/sortJob
Trace : Sort Count <2> | *****Sort Completed*****
Trace : Sort Count <2> | *****Initiating merge*****
Trace : Sort Count <2> | Sarting K Way Merge . File Size 678
Trace : Sort Count <1> | Creating file 674 complete
Trace : Sort Count <3> | Creating file 677 complete
Trace : Sort Count <1> | Creating file 675 complete
Trace : Sort Count <0> | Creating file 676 complete
Trace : Sort Count <3> | Creating file 678 complete
Trace : Sort Count <3> | *****Sort Completed*****
Trace : Sort Count <3> | *****Initiating merge*****
Trace : Sort Count <3> | Sarting K Way Merge . File Size 678
Trace : Sort Count <0> | Creating file 677 complete
Trace : Sort Count <1> | Creating file 676 complete
Trace : Sort Count <0> | Creating file 678 complete
Trace : Sort Count <0> | *****Sort Completed*****
Trace : Sort Count <0> | *****Initiating merge*****
Trace : Sort Count <0> | Sarting K Way Merge . File Size 678
Trace : Sort Count <1> | Creating file 677 complete
Trace : Sort Count <1> | Creating file 678 complete
Trace : Sort Count <1> | *****Sort Completed*****
Trace : Sort Count <1> | *****Initiating merge*****
Trace : Sort Count <1> | Sarting K Way Merge . File Size 678
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@1a3394c4
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@5cd52f64
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@1a5bd315
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@14f479f2
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@5212da8e
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@20f3e0a4
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@5f595e3a
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@562d4414
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@10cbfffb4
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@52b4a524
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@132fd0b5
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@6e4c272f
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@18644181
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@55c973c9
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@667463bb
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@535a5800
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@22f60351
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@64d68070
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@1e47eeff
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@2a4b8288
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@2baff501
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@35e68c86
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@587af718
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@73002ffa
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@36acf39a
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@4928f7c5
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@34fd5862
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@64788b8d
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@24d28295
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@76610884
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@543a0e6
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@3f76a175
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@489cae3a
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@4cb3f780
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@30ab7dfe
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@774e304b
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@31a855d1
```

128GB Output

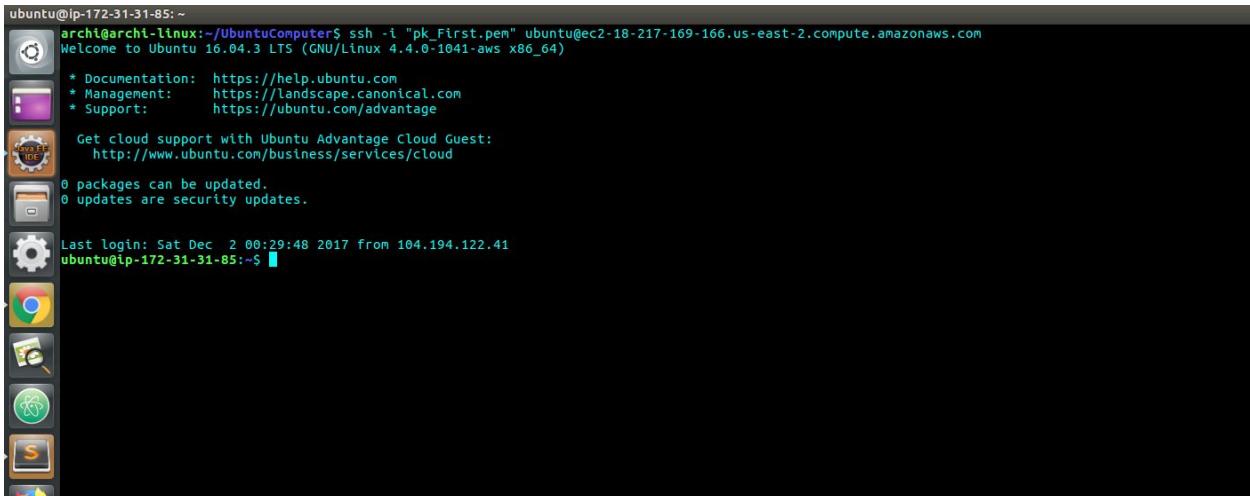
```
ubuntu@ip-172-31-4-166: ~          ubuntu@ip-172-31-31-85: ~          ubuntu@ip-
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@11000f8f
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@79fcaddb
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@683a02b1
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@3bccd6ab
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@5f595e3a
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@2edd67c6
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@298375d0
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@6c81fe84
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@52b4a524
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@177463aa
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@3fb069e5
Trace : Sort Count <3> | Completed K Way Merge
Trace : Sort Count <3> | *****Deleting temp file*****
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@6e4c272f
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@5da774f7
Trace : Sort Count <3> | *****merge completed*****
Trace : Sort Count <3> | Current count | 3
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@6c058b6e
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@2907088a
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@2a4b8288
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@6fb8b766
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@20f3e0a4
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@18644181
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@3f76a175
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@489cae3a
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@587af718
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@4928f7c5
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@cae01cb
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@7661d884
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@47669d59
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@38734ac7
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@314f08ff
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@5e9581f2
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@4cb3f780
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@437e1b39
Trace : Sort Count <1> | Completed K Way Merge
Trace : Sort Count <1> | *****Deleting temp file*****
Trace : Sort Count <0> | Completed K Way Merge
Trace : Sort Count <0> | *****Deleting temp file*****
Trace : Sort Count <0> | *****merge completed*****
Trace : Sort Count <0> | Current count | 2
Trace : Sort Count <1> | *****merge completed*****
Trace : Sort Count <1> | Current count | 1
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@74bc1ec1
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@8e0f9c4
Trace : Sort Count <2> | File Reached end java.io.BufferedReader@2a67af6d
Trace : Sort Count <2> | Completed K Way Merge
Trace : Sort Count <2> | *****Deleting temp file*****
Trace : Sort Count <2> | *****merge completed*****
Trace : Sort Count <2> | Current count | 0
File Added SORT_0MERGE
File Added SORT_1MERGE
File Added SORT_2MERGE
File Added SORT_3MERGE
File Size 4
Total Time taken in Seconds-----> 5760
*****Completed*****
=====Starting Sort Count :: 4=====
Trace : Sort Count <4> | Sarting K Way Merge . File Size 4
```

Val Sort Confirmation :



```
*****Completed*****
=====>Starting Sort Count :: 4=====
Trace : Sort Count <4> | Sarting K Way Merge . File Size 4
Trace : Sort Count <4> | File Reached end java.io.BufferedReader@3badb99e
Trace : Sort Count <4> | File Reached end java.io.BufferedReader@356a509b
Trace : Sort Count <4> | File Reached end java.io.BufferedReader@3155d2ee
Trace : Sort Count <4> | File Reached end java.io.BufferedReader@74575aa0
Trace : Sort Count <4> | Completed K Way Merge
ubuntu@ip-172-31-4-166:~$ ./valsort input/sortJob/SORT_4MERGE
Records: 1374389534
Checksum: 2a2898baf98ab2
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-4-166:~$
```

EC2 i3.x4 large instance login



```
ubuntu@ip-172-31-31-85:~ archi@archi-linux:~/UbuntuComputer$ ssh -i "pk_First.pem" ubuntu@ec2-18-217-169-166.us-east-2.compute.amazonaws.com
Welcome to Ubuntu 16.04.3 LTS (GNU/Linux 4.4.0-1041-aws x86_64)

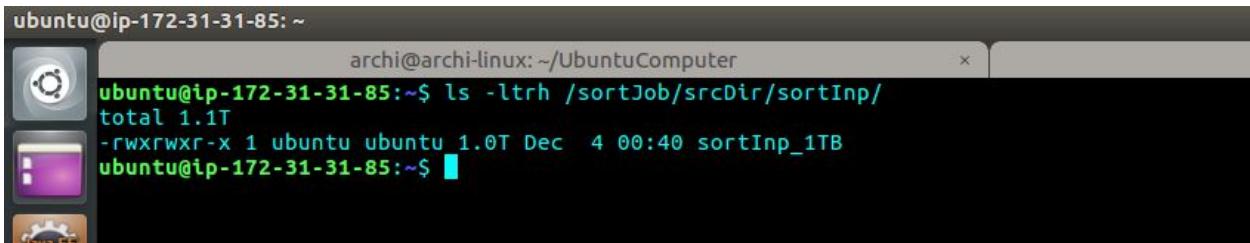
 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

 Get cloud support with Ubuntu Advantage Cloud Guest:
   http://www.ubuntu.com/business/services/cloud

 0 packages can be updated.
 0 updates are security updates.

Last login: Sat Dec  2 00:29:48 2017 from 104.194.122.41
ubuntu@ip-172-31-31-85:~$
```

1 TB data Generation

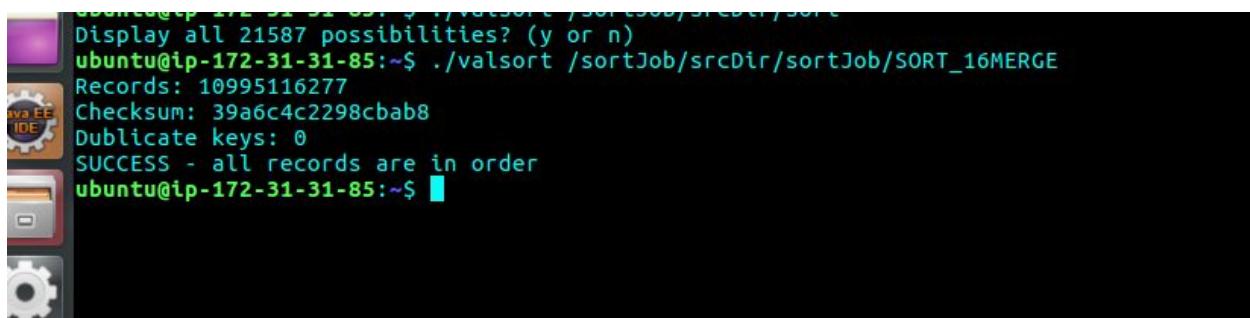


```
ubuntu@ip-172-31-31-85:~ archi@archi-linux:~/UbuntuComputer
ubuntu@ip-172-31-31-85:~$ ls -ltrh /sortJob/srcDir/sortInp/
total 1.1T
-rwxrwxr-x 1 ubuntu ubuntu 1.0T Dec  4 00:40 sortInp_1TB
ubuntu@ip-172-31-31-85:~$
```

1 TB final OP

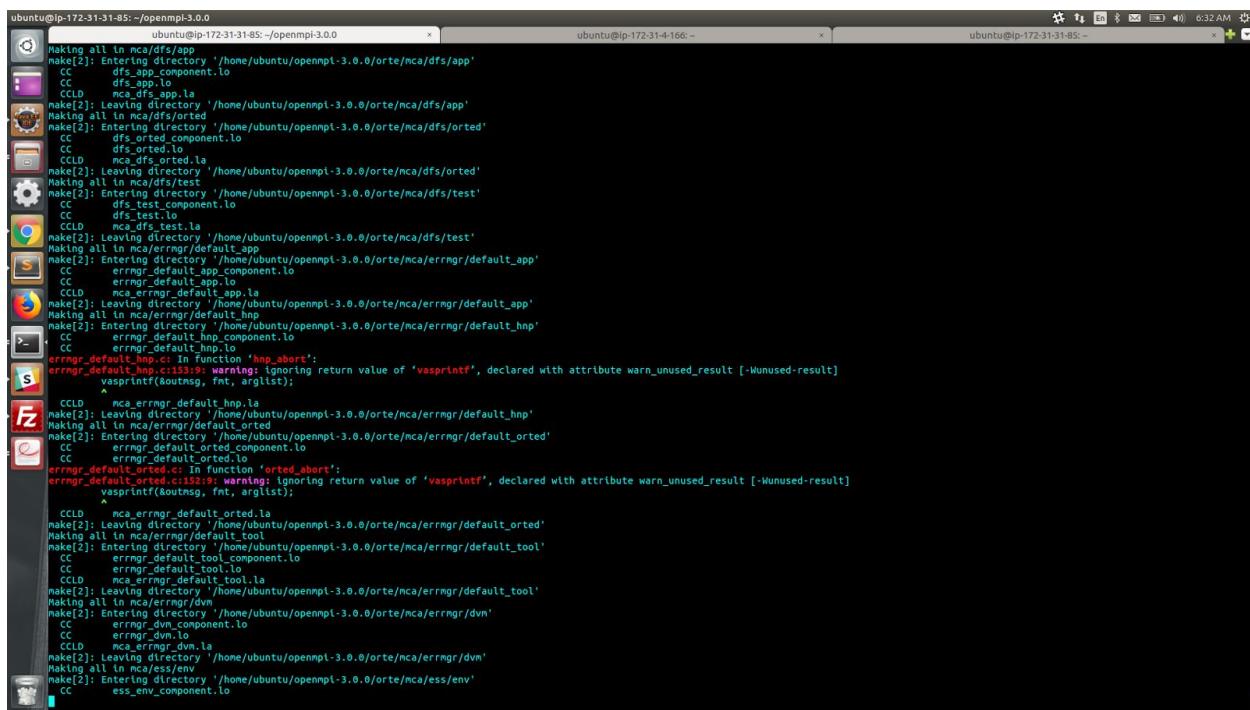
```
ubuntu@ip-172-31-31-85: ~
Trace : Sort Count <3> | File Reached end java.io.BufferedReader@79ca4be5
Trace : Sort Count <6> | Completed K Way Merge
Trace : Sort Count <6> | *****Deleting temp file*****
Trace : Sort Count <2> | Completed K Way Merge
Trace : Sort Count <2> | *****Deleting temp file*****
Trace : Sort Count <12> | Completed K Way Merge
Trace : Sort Count <12> | *****Deleting temp file*****
Trace : Sort Count <8> | Completed K Way Merge
Trace : Sort Count <8> | *****Deleting temp file*****
Trace : Sort Count <3> | Completed K Way Merge
Trace : Sort Count <3> | *****Deleting temp file*****
Trace : Sort Count <6> | *****merge completed*****
Trace : Sort Count <6> | Current count | 4
Trace : Sort Count <12> | *****merge completed*****
Trace : Sort Count <12> | Current count | 3
Trace : Sort Count <2> | *****merge completed*****
Trace : Sort Count <2> | Current count | 2
Trace : Sort Count <3> | *****merge completed*****
Trace : Sort Count <3> | Current count | 1
Trace : Sort Count <8> | *****merge completed*****
Trace : Sort Count <8> | Current count | 0
File Added SORT_8MERGE
File Added SORT_9MERGE
File Added SORT_4MERGE
File Added SORT_5MERGE
File Added SORT_1MERGE
File Added SORT_10MERGE
File Added SORT_12MERGE
File Added SORT_11MERGE
File Added SORT_13MERGE
File Added SORT_14MERGE
File Added SORT_2MERGE
File Added SORT_3MERGE
File Added SORT_7MERGE
File Added SORT_6MERGE
File Added SORT_0MERGE
File Added SORT_15MERGE
File Size 16
Total Time taken in Seconds-----> 32760
*****Completed*****
=====Starting Sort Count :: 16=====
Trace : Sort Count <16> | Sarting K Way Merge . File Size 16
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@6e9ef235
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@60591196
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@64669945
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@21641bad
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@38724aee
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@100f39d7
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@40f33622
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@414be4e2
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@6bc4e31f
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@2ecae31
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@5bf7da68
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@45055900
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@99596a4
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@6eef2c81
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@bae8fae
Trace : Sort Count <16> | File Reached end java.io.BufferedReader@640c0a82
Trace : Sort Count <16> | Completed K Way Merge
ubuntu@ip-172-31-31-85:~$
```

1 TB Valsort Confirmation



```
ubuntu@ip-172-31-31-85:~/valsor...$ ./valsor... /sortJob/srcDir/sortJob/SORT_16MERGE
Display all 21587 possibilities? (y or n)
ubuntu@ip-172-31-31-85:~/~$ ./valsor... /sortJob/srcDir/sortJob/SORT_16MERGE
Records: 10995116277
Checksum: 39a6c4c2298cbab8
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-31-85:~/~$
```

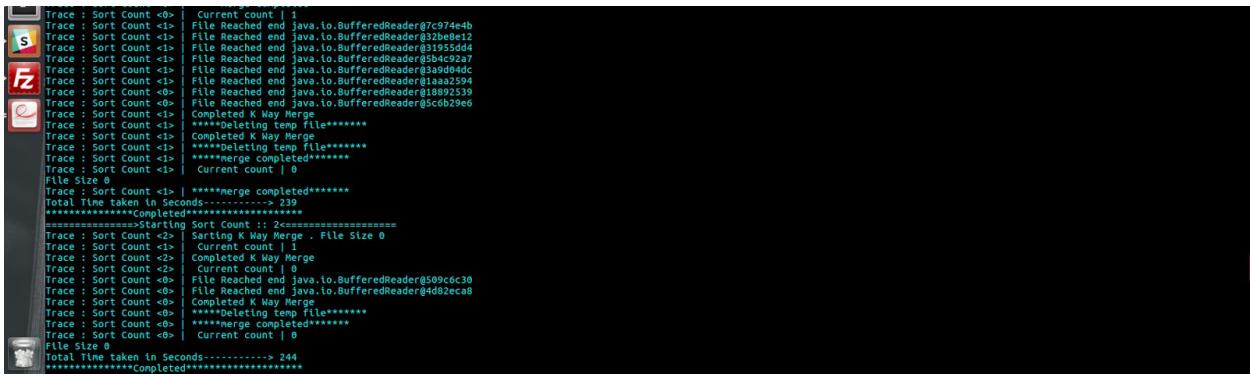
MPI Install Screen shot



```
ubuntu@ip-172-31-31-85:~/openmpi-3.0.0
  ubuntu@ip-172-31-31-85:~/openmpi-3.0.0
  ubuntu@ip-172-31-31-85:~
```

The screenshot shows three terminal windows side-by-side. The leftmost window displays the compilation log for the OpenMPI 3.0.0 source code, specifically the mca/dfs/app directory. It shows the 'make' command entering the directory, compiling components like 'dfs_app.lo' and 'dfs_orted.lo', and linking them into 'mca_dfs_orted.la'. It also shows the compilation of 'dfs_test_component.lo' and 'mca_dfs_test.lo'. The middle window shows the 'make' command entering the 'mca/errmgr/default_app' directory, which contains 'errmgr_default_app_component.lo' and 'errmgr_default_app.lo'. The rightmost window shows the 'make' command entering the 'mca/errmgr/default_orted' directory, containing 'errmgr_default_orted_component.lo' and 'errmgr_default_orted.lo'. All three windows show the standard Linux terminal interface with a black background and white text.

MPI Local Run screenshot on 2 GB data



```
Trace : Sort Count <0> | Current count | 1
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@7c974e4b
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@32be8e12
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@31955dd4
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@3b4c9287
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@1aa2594
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@18892539
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@5ccb29e6
Trace : Sort Count <1> | Completed K Way Merge
Trace : Sort Count <1> | Starting K Way Merge . File Size 0
Trace : Sort Count <1> | Current count | 1
Trace : Sort Count <1> | Completed K Way Merge
Trace : Sort Count <1> | Current count | 0
Trace : Sort Count <1> | ****Deleting temp file*****
Trace : Sort Count <1> | ****Merge completed*****
Trace : Sort Count <1> | Current count | 0
File Size 0
Trace : Sort Count <1> | ****Merge completed*****
Total Time taken in Seconds-----> 239
*****Completed*****
```



```
=====Starting Sort Count :: 2=====
Trace : Sort Count <2> | Starting K Way Merge . File Size 0
Trace : Sort Count <1> | Current count | 1
Trace : Sort Count <1> | Completed K Way Merge
Trace : Sort Count <1> | Current count | 0
Trace : Sort Count <1> | File Reached end java.io.BufferedReader@509c6c30
Trace : Sort Count <0> | File Reached end java.io.BufferedReader@4d82eca8
Trace : Sort Count <0> | Completed K Way Merge
Trace : Sort Count <0> | ****Deleting temp file*****
Trace : Sort Count <0> | ****Merge completed*****
Trace : Sort Count <0> | Current count | 0
File Size 0
Total Time taken in Seconds-----> 244
*****Completed*****
```